

**SF-83-1 SUPPORTING STATEMENT
(Part A)**

for the

2023

Survey of Doctorate Recipients

TABLE OF CONTENTS

A. JUSTIFICATION.....	1
1. NECESSITY FOR INFORMATION COLLECTION.....	2
2. PURPOSE AND USE OF THE INFORMATION.....	3
3. USE OF AUTOMATED, ELECTRONIC, MECHANICAL, OR OTHER TECHNOLOGICAL TECHNIQUES.....	6
4. EFFORTS TO IDENTIFY DUPLICATION.....	7
5. IMPACT ON SMALL BUSINESS.....	8
6. CONSEQUENCES OF LESS FREQUENT DATA COLLECTION.....	8
7. SPECIAL CIRCUMSTANCES.....	8
8. CONSULTATIONS OUTSIDE THE AGENCY.....	8
9. PAYMENT OR GIFTS TO RESPONDENTS.....	10
10. ASSURANCE OF CONFIDENTIALITY.....	12
11. JUSTIFICATION FOR ASKING SENSITIVE QUESTIONS.....	13
12. BURDEN ESTIMATE.....	13
13. COST TO RESPONDENTS.....	14
14. COST TO THE FEDERAL GOVERNMENT.....	14
15. CHANGES FROM PRIOR CYCLE.....	14
16. PLAN FOR TABULATION OR PUBLICATION.....	14
17. DISPLAY OF OMB EXPIRATION DATE.....	15
18. EXCEPTION TO THE CERTIFICATION STATEMENT.....	15

LIST OF APPENDICES

Appendix A – NSF Act of 1950; America COMPETES Reauthorization Act of 2010.....	A-1
Appendix B – First Federal Register Announcement.....	B-1
Appendix C – Examples of SDR Data in Use by the Research Community.....	C-1
Appendix D.1 – SDR 2023 Questionnaire Changes.....	D.1-1
Appendix D.2 – 2021 SDR Questionnaire.....	D.2-1
Appendix D.3 – NSCG/SDR Retirement Module Development Literature Review.....	D.3-1
Appendix D.4 – Retirement Module Phase 2 Cognitive Testing Final Report.....	D.4-1
Appendix D.5 – Survey of Doctorate Recipients Sexual Orientation and Gender Identity Question Experiment Plan.....	D.5-1
Appendix D.6 – Survey of Doctorate Recipients Race and Ethnicity Question Experiment Plan.....	D.6-1
Appendix E.1 – Draft 2023 SDR Survey Mailing Materials.....	E.1-1
Appendix E.2 – Prefield Contact Materials for 2025 SDR.....	E.2-1
Appendix F – 2023 SDR Sample Allocation and Selection Table.....	F-1

2023 SURVEY OF DOCTORATE RECIPIENTS SUPPORTING STATEMENT

A. JUSTIFICATION

This request is for a three-year renewal of the previously approved OMB clearance for the Survey of Doctorate Recipients (SDR). The SDR has historically served as a valuable source of information on U.S.-trained science, engineering, and health doctorate recipients. The SDR was last conducted in 2021. The OMB clearance for the 2021 SDR expires July 31, 2024 (OMB No 3145-0020). There were no terms of clearance from OMB for the previously approved OMB clearance.

The sample size of 125,246 in the 2023 cycle is slightly less than the prior round of 125,938 cases. The 2023 sample size reflects a decision to exclude chronic nonrespondents from all eligible returning sample members from the 2021 cycle along with including 10,000 newly sampled graduates. The decision to exclude chronic nonrespondents from all eligible returning sample cases from the prior cycle is to improve the efficiency of the SDR data collection as well as reduce the burden on sample members who have shown that they have no interest in responding to the SDR or have never been located. Two groups of cases were identified as chronic nonrespondents for the 2023 cycle:

- 2017 SDR new cohort cases that have never responded to the SDR (i.e., did not respond in 2017, 2019, and 2021); and
- 2015 SDR supplemental sample cases, added in the 2019 cycle, that have never responded to the SDR (i.e., did not respond in 2019 and 2021).

In addition, NCSES is proposing the following changes to the 2023 SDR operations:

- 1) **Inclusion of a new retirement module.** Though the SDR has consistently included questions about the timing of retirement, it fails to capture information about retirement experiences or details about what can be a complex process from the workforce to retirement. For many full-time employed individuals, retirement is a process with one or more stops along the way – usually with another employer, sometimes in another part of the country, and often part-time. Currently, there is limited data that provides insight on an individual’s pathway toward full retirement. The aim of this new module is to collect objective data on how individuals experience retirement including, but not limited to, partial and phased retirement approaches, post-retirement returns to paid employment, and permanent departures from the workforce. The new questions also aim to capture some of the circumstances and reasons for retirement, as well as identifying various post-employment activities. The new module was initially developed in 2022 and received expert review in spring 2022. The resulting module subsequently underwent three rounds of cognitive testing in fall-winter 2022-2023. For more information on this research, see Appendices D.3 and D.4.
- 2) **Removal of questions that asked about the effect of the coronavirus pandemic** on sample members' employment situation in 2021. Items specific to the pandemic will be removed. However, due to the prevalence of remote work, **the subset of items about telework will be updated and included for 2023.**
- 3) **Implementation of an incentive experiment.** NCSES proposes to include an incentive experiment in 2023 to assess the use of post-paid electronic gift cards (a method new to SDR) compared to personalized checks mailed with a request for survey participation (a traditional SDR incentive method). See Section 9 below and Section 3.4 in Supporting Statement Part B for more detail.

- 4) **Implementation of an experiment on including questions about sexual orientation and gender identity.** NCSES will continue its research on how to collect sexual orientation and gender identity (SOGI) data from doctorate recipients that will yield accurate and usable data while ensuring privacy and confidentiality protections. An experiment will test three batteries of questions for measuring biological sex, gender identity, and sexual orientation. The experimental design is described in Appendix D5.
- 5) **Implementation of an experiment on including new ways to ask questions about race and ethnicity.** NCSES intends to test new race and ethnicity questions that combine the questions on race and ethnicity into a single question with disaggregation by detailed categories. The information will support the ongoing research led by the Office of Management and Budget (OMB) in revising OMB's 1997 Statistical Policy Directive No. 15: Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity. The experimental design for this effort is described in Appendix D6. In order to produce official statistics from the SDR race and ethnicity, as in past SDR years, race and ethnicity for each respondent will be derived from their SED responses. For those who did not provide race and ethnicity in the SED, or in subsequent years of the SDR, their race and ethnicity will be imputed. This last step differs from past SDR years, in which respondents who did not provide their race and ethnicity in the SED or previous SDR years were asked the race and ethnicity questions again in the current SDR web and computer-assisted telephone interviews response modes.

Further details about proposed changes to the questionnaire are in Section 4 of Supporting Statement Part B and Appendix D.1.

1. NECESSITY FOR INFORMATION COLLECTION

Established within the NSF by the America COMPETES Reauthorization Act of 2010 § 505, codified in the National Science Foundation Act of 1950, as amended, the National Center for Science and Engineering Statistics (NCSES) serves as a central Federal clearinghouse for the collection, interpretation, analysis, and dissemination of objective data on science, engineering, technology, and research and development for use by practitioners, researchers, policymakers, and the public. Information obtained through the SDR is critically important to NCSES's ability to measure the education and employment outcomes of scientists and engineers. Furthermore, the SDR and NCSES's National Survey of College Graduates (NSCG) are coordinated in both survey content and timing to form data collections that serve as the nation's only source of comprehensive information about the size and characteristics of the science and engineering (S&E) workforce.¹ These data are solicited under the authority of the NSF Act of 1950², as amended, and are central to the analysis presented in a pair of congressionally mandated reports^{3,4} published by NSF:

- *Science and Engineering Indicators*
- *Diversity and STEM: Women, Minorities, and Persons with Disabilities.*

In addition, the Science and Engineering Equal Opportunities Act of 1980 directs NSF to provide to Congress and the Executive Branch an "accounting and comparison by sex, race, and ethnic group and by

¹ The S&E workforce includes individuals with degrees or occupations in computer and mathematical sciences, life sciences, physical sciences, social sciences, engineering, and health sciences.

² See Appendix A.

³ 42 U.S. Code § 1863(j)(1)

⁴ 42 U.S. Code § 1885(a), 1885(d)

discipline, of the participation of women and men in scientific and engineering positions.”⁵ The SDR and NSCG provide much of the information to meet this mandate. The coordinated designs of these two surveys were developed throughout the past two decades and are based on recommendations from the National Research Council’s Committee on National Statistics (CNSTAT) report to NSF.⁶

2. PURPOSE AND USE OF THE INFORMATION

SDR Background

The SDR provides education and employment-related information on scientists and engineers who were awarded a research doctoral degree from a U.S. institution in a science, engineering, or health (SEH) field. A research doctorate is a doctoral degree that (1) requires completion of an original intellectual contribution in the form of a dissertation or an equivalent culminating project (e.g., a published manuscript) and (2) is not primarily intended as a degree for the practice of a profession.

The 2023 SDR is comprised of two sample components:

- A continuing panel of doctorate recipients from the 2021 survey cycle who remain eligible and are under 76 years of age.
- A new cohort component that adds newly eligible doctorate recipients from academic years 2020 and 2021, also under 76 years of age.

The continuing panel portion of the SDR sample provides information on the cohorts of doctorate recipients who earned their degrees prior to July 2019. The new cohort sample from the two most recent doctorate award years provides important data on the early career experiences of new doctorate recipients with SEH degrees.

Since 2010, the SDR had included an international component of U.S.-trained doctorate recipients who received their degrees as of 2001 (i.e., 21st century graduates only). The more recently redesigned 2015 SDR cycle used the cumulative Doctorate Records File, the ongoing SDR sampling frame, to refresh the 2015 SDR sample and significantly increase its size and target population representation. The redrawn sample expanded the international component to be representative of all academic years dating back to 1961. The 2023 SDR will continue to field a sample of SEH doctorate holders predicted to reside either in or outside of the U.S. on the survey reference date of 1 February 2023. Based on the data from the 2020 and 2021 Survey of Earned Doctorates (SED), which is the frame for the 2023 SDR new cohort, 36% of the SEH doctorates were awarded to temporary visa holders and 23% of temporary visa holders planned to leave the U.S. upon graduation.⁷ Thus, the 2023 SDR will yield information about the educational, employment, and demographic characteristics of U.S.-trained SEH doctorate recipients living and working both in the U.S. and abroad.

The 2015 SDR sample size expansion from approximately 45,000 to 120,000 was also designed to support estimation of employment outcomes for a larger number of fine fields of degree (FFOD) captured in the SED. This large sample size will continue to be maintained for the 2023 SDR cycle. A total of 115,246 of 125,938 sampled cases in the 2021 survey remain eligible for the 2023 cycle and will be increased by 10,000 new sample members from the 2020 and 2021 SED cohorts for a total sample size of 125,246 cases.

⁵ 42 U.S. Code § 1885(d)

⁶ National Research Council, Committee on National Statistics. 1989. *Surveying the Nation’s Scientists and Engineers: A Data System for the 1990s*. Washington: National Academy Press.

⁷ The SED gathers information yearly from all new research doctorates awarded by U.S. institutions. Detailed information about the SED can be found at <http://www.nsf.gov/statistics/srvydoctorates/>.

Originally, the SDR was designed to produce employment outcome estimates for various analytical domains defined by broad aggregated fields of degree and demographic characteristics. The objective of the 2015 sample expansion and the more recent 2019 sample design modification was to meet new FFOD estimation goals while maintaining the estimation goals for the traditional (historic) key analytic domains.

In 2019, NCSSES modified its fine field of degree stratification variable based on results from 2015 and 2017 estimation capabilities. The modified detailed field of study sampling stratification variable for the SDR better aligns with the NCSSES Taxonomy of Disciplines (TOD). The NCSSES TOD more closely aligns with the National Center for Education Statistics (NCES) Classification of Instructional Programs (CIP). This revised stratification supports a more stable sampling design and more reliable estimation capability in subsequent cycles of the SDR. As with the 2021 SDR, the 2023 SDR stratifies the sample frame by 77 detailed fields rather than the more than 200 FFODs used in the 2015 and 2017 SDR cycles.

At the conclusion of the 2023 SDR survey cycle, all sampled respondents will be used to develop cross-sectional estimates describing the U.S.-trained SEH doctorate recipient population. In addition, the respondents from the 40,000-case sample identified in the 2019 SDR as the SDR longitudinal sample will be the basis for deriving longitudinal estimates from these data. The 40,000 longitudinal sample cases represent the 2015 SDR target population moving forward and will be weighted and maintained through the 2025 cycle of the SDR to provide longitudinal data for the 10-year period 2015 to 2025. The 2023 SDR longitudinal sample will provide longitudinal data over the 8-year period of 2015-2023.

The SDR enables government agencies to assess the scientific and engineering resources available in the U.S. to business, industry, and academia, and provide a basis for the formulation of the nation's S&E workforce policies. For example, educational institutions can use the SDR data in establishing and modifying scientific and technical curricula, while various industries can use the information to develop recruitment and remuneration policies.

Policymakers, researchers, and others use information from the SDR and NSCG to answer questions about the number, employment, education, and characteristics of the S&E workforce. Because the SDR and NSCG provide up-to-date and nationally representative data, policymakers and researchers use these datasets to address questions on topics such as: employment of foreign-born or foreign-degreed scientists and engineers, the transition from higher education to the workforce, the role and importance of postdocs as research personnel, diversity in both education and employment, the implications of an aging cohort of scientists and engineers as baby boomers reach retirement age, and long-term trends in the S&E workforce.

Findings from the 2023 SDR will enable NCSSES to continue reporting employment patterns of recent SEH doctorate recipients, as well as about the more experienced doctorate recipients in the labor market. As noted, the sample size and composition enable NCSSES to produce reliable estimates of employment outcomes by at least 77 detailed fields of degrees listed in the NCSSES TOD. The SDR data are made available through published reports, the Scientists and Engineers Statistical Data System (SESTAT), public use data files, and licenses for restricted-use data files. The online data tool, available at <https://ncesdata.nsf.gov/sestat>, allows users to create customized data tabulations in a user-specified subject area using the public use file. In addition, the SDR public-use files are available for download through the NCSSES data downloads web page at <https://ncesdata.nsf.gov/datadownload>.

In addition, the 2023 survey cycle will continue with the longitudinal component of the SDR target population and its sample first defined and selected in the 2019 data collection. The 40,000-case longitudinal sample provides the ability to derive longitudinal estimates from the 2015 SDR target population on a variety of policy-relevant topics. SDR longitudinal data can be used to address important

issues such as changes in employer, occupation, field of research, international mobility, income over time and retirement. Within the academic sector, SDR longitudinal data can address issues such as length of post-doctoral training, achieving tenure, and changes in location especially for married SEH doctorate-holding couples. As noted previously, the longitudinal sample will represent employment and other outcomes for the 2015 SDR population through the year 2025 based on 6 cycles of the SDR data.

Uses for Policy Discussion

Data from NCSES's surveys are used in policy discussions of the executive and legislative branches of Government, the National Science Board, NSF management, the National Academy of Sciences, Engineering, and Medicine, professional associations, and other private and public organizations. Some recent examples of the current use of the SDR data, and the combined SDR and NSCG data, are as follows:

- The National Science Board (NSB) used the combined SDR and NSCG data in its investigation to develop national policies for the S&E workforce⁸;
- The Committee for Equal Opportunity in Science and Engineering (CEOSE), an advisory committee to NSF and other government agencies, established under 42 U.S.C. §1885c, has been charged by the U.S. Congress with advising NSF in assuring that all individuals are empowered and enabled to participate fully in science, mathematics, engineering and technology. Every two years CEOSE prepares a congressionally mandated report that makes extensive use of the SDR and NSCG data to highlight key areas of concerns relating to students, educators, and technical professionals;
- A recent publication, *The Long-Term Stay Rates of International STEM PhD Graduates* used the SDR data to examine how many international STEM PhD graduates choose to stay in the U.S. after earning their degrees⁹;
- The importance of information on the S&E workforce to inform public policy can be seen in discussions of the NSB's Task Group on Science, Technology, Engineering, and Math (STEM) Innovators. The task group used the SDR and NSCG data to inform its deliberations about the S&E workforce and these data were part of the final report¹⁰; and
- SDR data is frequently used to better understand issues surrounding diversity and inclusion as demonstrated by these articles from 2022: *Preparing Industry Leaders: The Role of Doctoral Education and Early Career Management Training in the Leadership Trajectories of Women STEM PhDs*¹¹, *Disability disparities in STEM: Gaps in salaries and representation for doctorate recipients with disabilities in the U.S., 2019*¹², and *Gender, Race/Ethnicity and Research Funding*.¹³

Uses by NSF and NCSES

⁸ <https://www.nsf.gov/pubs/2018/nsb20187/nsb20187.pdf>, <http://nsf.gov/nsb/publications/2015/nsb201510.pdf> and <http://www.nsf.gov/nsb/documents/2003/nsb0369/nsb0369.pdf>

⁹ <https://cset.georgetown.edu/wp-content/uploads/CSET-The-Long-Term-Stay-Rates-of-International-STEM-PhD-Graduates.pdf>

¹⁰ <http://www.nsf.gov/nsb/publications/2010/nsb1033.pdf>

¹¹ <https://link.springer.com/article/10.1007/s11162-021-09655-7#citeas>

¹² <https://www.medrxiv.org/content/10.1101/2022.12.04.22283081v2>

¹³ <https://www.nber.org/sites/default/files/2022-05/GenderRaceResearchFundingv3.pdf>

The SDR data were used extensively in the latest versions of the congressionally mandated biennial reports: *Science and Engineering Indicators, 2022* and *Diversity and STEM: Women, Minorities and Persons with Disabilities, 2023*.

In addition, NCSES used the SDR data and the combined SDR and NSCG data in recent reports such as:

- *U.S. Residing Doctoral Scientists and Engineers Report Modest Professional Impacts from the Coronavirus Pandemic: Findings from the 2021 Survey of Doctorate Recipients*, NSF 23-318, January 27, 2023.
- *Science and Engineering State Profiles*, updated October 2022.
- *Labor Force Transitions of U.S.-Trained Doctoral Scientists and Engineers: Findings from a New Longitudinal Panel*, NSF 22-327, June 1, 2022.
- *U.S. Employment Higher in the Private Sector than in the Education Sector for U.S.-Trained Doctoral Scientists and Engineers: Findings from the 2019 Survey of Doctorate Recipients*, NSF 21-319, April 21, 2021.
- *Where Are They Now? Most Early Career U.S.-Trained S&E Doctorate Recipients with Temporary Visas at Graduation Stay and Work in the United States after Graduation*, NSF 21-336, August 19, 2021.
- *Number of Women with U.S. Doctorates in Science, Engineering, or Health Employed in the United States More Than Doubles since 1997*, NSF 19-307, February 12, 2019.
- Data Tables: *Survey of Doctorate Recipients, 2021, January 2023*.
- Data Tables: *Survey of Doctorate Recipients, Longitudinal Data: 2015-19*, June 2022.
- Data Tables: *Survey of Doctorate Recipients, 2019*, April 2021.
- Data Tables: *Survey of Doctorate Recipients, 2017*, February 2019.

All NCSES publications can be accessed on the NCSES website at <http://www.nsf.gov/statistics>.

The SDR and the combined SDR and NSCG data are presented at conferences and professional meetings by NCSES staff and survey contractor staff, such as the annual meeting of the Association for Institutional Research, the American Association for Public Opinion Research, the American Educational Research Association, and the Joint Statistical Meetings. See Appendix C for examples of these presentations.

Uses by Researchers and Analysts

Since 2009, NCSES has distributed more than 6,500 copies of SDR public-use files (2003, 2006, 2008, 2010, 2013, 2015, 2017, and 2019 survey cycles), as well as over 6,945 copies of the combined SDR and NSCG data's public-use files (1993-2013 survey cycles). There are currently 48 active SDR restricted-use licenses. Select recent citations by researchers using the SDR data are provided in Appendix C.

3. USE OF AUTOMATED, ELECTRONIC, MECHANICAL, OR OTHER TECHNOLOGICAL TECHNIQUES

As with earlier cycles, the 2023 SDR will collect data using three modes:

- Self-administered online (or web) surveys (including access from mobile devices);
- Paper self-administered questionnaires (mail); and

- Computer-assisted telephone interviews (CATI).

The data collection effort will emphasize response by web, use a mobile optimized web instrument, incorporate adaptive design techniques, and store contact history information in a case management system. These systems are described below.

Participation via the online response option has increased steadily since the 2006 cycle, from 47% in 2006 to 93% in 2019 to almost 99% in 2021. In 2021, in addition to the online response option, 0.6% responded by paper questionnaire and 0.6% responded through CATI. Analysis indicates that the computer-assisted web and CATI modes result in more complete and concise survey data and follow-up contacting information compared to the mail questionnaire mode. As with the 2021 survey protocol, for 2023, nearly all sample members initially will be asked to participate via the web form with exception for those known to have vision problems, no internet access, or other circumstances that warrant their participation by other modes. Mailed letters will include the SDR URL and the sample members Personal Identification Number (PIN). Email invitations will include a live link to the starting page of the 2023 SDR. The online survey will be 508 compliant.

As in 2021, the 2023 online survey will be configured for use on mobile devices (e.g., smartphones and tablets) to ensure that the respondent experience is optimized regardless of the screen size or browser used to access the survey. The 2021 SDR obtained about 16% of the online responses via submission by a mobile device, up about five percentage points from 2019 (11%), and we expect a similar, if not higher, percentage of online submissions via a mobile device in 2023.

The 2023 SDR data collection effort will continue to use a comprehensive case management system (CMS) to track data captured across the three modes (web, mail, CATI) with additional features added to the system. The CMS is structured around a federated case management architecture design which is comprised of numerous components, user interface modules, and data sources that maintains records for all incoming and outgoing contacts with sample members, as well as complete address history information. The additional features support an adaptive design data collection strategy for quickly prioritizing cases and real-time response rate calculations for Primary Analytic Domains (PADs) such as location of residency, employment sector and demographic characteristics of the sampled cases. Additionally, the CMS will include methods to monitor incentive use by response status, which will allow follow-up contacts tailored to incentive receipt and usage.

4. EFFORTS TO IDENTIFY DUPLICATION

Overlap exists between the SDR and the NSCG in their target population and in survey content. The NSCG and the SDR both capture estimates of the roughly 1 million U.S.-trained SEH doctorates residing in the U.S., though the SDR also covers an estimated 162,000 U.S.-trained SEH doctorates residing outside of the U.S. and the NSCG covers an estimated 165,000 employed foreign-trained SEH doctorates residing in the United States. However, the SDR is designed to estimate the population of scientists and engineers that receive their research doctorate degrees in the U.S. with a greater level of precision than the NSCG. With this greater level of precision in mind, the SDR collects detailed information from U.S.-trained SEH doctorate recipients working in post-secondary academic institutions including their academic position, faculty rank, tenure status, and reasons for taking a postdoc if in a postdoctoral position. In contrast, and unlike the SDR, the NSCG collects information on all college graduates at the bachelor level and higher regarding veteran status, attainment of certifications and licenses, financial support for education, and community college enrollment history and does not sample an adequate number of SEH doctorate holders to allow for granular analysis of the most educated population of scientists and engineers as the SDR does.

Based on the 2021 SDR sample, approximately 400 individuals may be selected for both the 2023 NSCG and the 2023 SDR. Due to the content differences between the surveys, the relatively small number of expected duplicates, and the operational challenges of the deduplication process, NCSES will not de-duplicate individuals selected for participation in both the SDR and NSCG.

5. IMPACT ON SMALL BUSINESS

Not applicable. The SDR collects information from individuals only.

6. CONSEQUENCES OF LESS FREQUENT DATA COLLECTION

The SDR data are central to the analysis presented in a pair of congressionally mandated reports - *Science and Engineering Indicators* and *Diversity and STEM: Women, Minorities, and Persons with Disabilities*. Since these reports are published on a biennial schedule, they rely on the availability of updated data on the S&E workforce every two years. Conducting the SDR on a less frequent basis would prohibit NCSES from meeting its congressional mandate to produce these reports that contain an accurate accounting and comparison, by sex, race, and ethnic group and by discipline, of the participation of women and men in scientific and engineering occupations. The impact of not being able to meet this congressional mandate is that government, business, industry, and universities would have less recent data to use as a basis for formulating the nation's S&E policies.

Less frequent data collection would also impact the SDR data quality and value. Follow-up surveys every two years enable both time-series and longitudinal analyses of changes in the S&E workforce especially with respect to employment in and out of S&E occupations and the growth of these occupations.

Because the SDR has been coordinated with the NSCG on a biennial basis since 1993 to ensure the availability of current national S&E workforce data, degradation of either component jeopardizes their integrity and value. These combined surveys provide comprehensive information on the S&E workforce. Conducting either survey less frequently would make it more difficult and costly to locate sampled persons given their mobility, and the impact is likely to be higher attrition rate, higher potential for nonresponse bias, and less reliable cross-sectional and longitudinal estimates.

7. SPECIAL CIRCUMSTANCES

Not applicable. This data collection does not require any one of the reporting requirements listed.

8. CONSULTATIONS OUTSIDE THE AGENCY

Federal Register Announcement

In response to the Federal Register Notice for the SDR that appeared on December 6, 2022 [FR Doc. 2022-26422 Filed 2-5-22, Vol 87, Number 233], one public comment was received by the closing date of February 6, 2023. See Appendix B for the announcement.

On 6 December 2022, Dr. Andrew Reamer of George Washington University sent an email to NSF on behalf of the American Economic Association and the Industry Studies Association. He requested the draft information collection request (ICR) materials for the 2023 SDR. NSF responded to Dr. Reamer on 22 December 2022, explaining that the 2023 SDR ICR materials were in the process of being prepared and that there were no substantive changes planned, except that all the COVID-related items will be removed from the questionnaire and items pertaining to retirement will be added. He was

directed to past cycle SDR questionnaires on the NSF website, which would be updated to reflect the survey year.

Consultations Outside the Agency

NCSES routinely seeks the advice and guidance of survey methodologists, statisticians, demographers, researchers, data analysts, and policymakers to examine numerous issues related to further development and continuous quality improvement of the SDR. As noted in the 2017 and 2019 SDR OMB Information Collection Requests, NCSES previously convened a Human Resources Experts Panel (HREP) meeting in October 2016 to review the redesigned 2015 SDR sample expansion with a large number of data users and stakeholders, and to collect information on their research interests that could be addressed with the SDR longitudinal data. This effort was developed and pursuant to the OMB approved ICR for the 2017 SDR which stated the following terms:

Approved consistent with the understanding that the primary goal of the design is to allow cross-sectional reporting for fine fields; a secondary goal is to build an efficient, effective, and sustainable longitudinal design.

In response to these terms and in alignment with NCSES's mission, NCSES conducted research with the goal of developing a sustainable design that will maintain the integrity of the survey's cross-sectional estimates while producing reliable data to support longitudinal analysis.

2019 SDR Survey Sample Redesign

As noted in Section 2 (Background of the SDR), with consultation from statistical experts outside of the Agency, the 2019 SDR survey sample was redesigned to produce detailed estimates on at least 77 fine fields of study according to demographic and employment characteristics. The redesigned sample stratification allowed for the continued cross-sectional reporting of historical SDR timeseries estimates of employment outcomes while also producing more detailed outcomes on fine fields of degree in subsequent cycle of data collection.

In addition, NCSES worked with the 2019 SDR survey contractor to conduct research on fielding prefilled web and CATI forms to facilitate an efficient and effective longitudinal design. Findings from the *2020 SDR Dependent Interviewing Survey Pilot Study* were used to modify dependent interviewing protocols for the 2023 SDR web and CATI instrument.

Federal Committee on Statistical Methodology Research Conference (FCSM) presentation: (March 9, 2018)

In addition to the changes in sample size noted above, NCSES had promoted the development of the SDR longitudinal design by presenting at the 2018 FCSM Research Conference. The presentation focused on the complexities of balancing cross-sectional and longitudinal goals while building an effective, efficient, and sustainable longitudinal design into the newly expanded cross-sectional SDR sample. Specifically, the presentation noted that to enhance SDR's utility and meet dual cross-sectional and longitudinal goals, longitudinal panels within the refreshed sample need to be established formally and maintained over time. The presentation also highlighted NCSES's on-going research efforts described above. The presentation concluded by noting that the process of developing a longitudinal design to enhance overall survey utility is a good application of a major survey redesign effort and the approaches and lessons learned from the SDR experience could be informative for other researchers involved in survey design.

Development of Retirement-Related Survey Items

NCSES has undertaken research to develop a more comprehensive survey item module on retirement in 2022 and 2023 (see Appendices D.3 and D.4). The new module will be included in the 2023 SDR and is in response to stakeholder feedback and a Consensus Study Report from the National Academies of Sciences, Engineering, and Medicine issued in 2022, *Understanding the Aging Workforce: Defining a Research Agenda*¹⁴, which suggests more research is needed to better understand, “..the micro-level impacts on older adults’ work and retirement of large-scale social changes—in technology, the economy, the labor market, and society writ large...Beneficial future research could explore contemporary—and changing—experiences of work and retirement and the conditions shaping health and well-being.”

9. PAYMENT OR GIFTS TO RESPONDENTS

Incentive use for the SDR began during the 2003 survey cycle and has continued for all subsequent cycles. The 2023 SDR incentive plan is modeled after the effective approach used in SDR survey cycles since 2013, using an early-stage incentive protocol for sample members identified as reluctant to respond and a late-stage incentive protocol to increase survey response and mitigate nonresponse bias of underrepresented subgroups. The incentive amount is set to be \$30 based on the findings of previous SDR incentive experiments and used for survey cycles during 2006 – 2019. In SDR 2021, an experiment varying the incentive amount offered (i.e., no incentive, \$10, or \$30) was conducted and the results showed significant declines in response rate from groups receiving incentive lower than \$30. Therefore, the incentive amount is reverted to \$30 for 2023 SDR.

The incentive-usage objectives are to:

- Minimizing differential unit nonresponse for key domains while maximizing survey unit and item response;
- Decrease the number of sample member contacts; and
- Expedite response from reluctant sample members.

Incentives will only be sent or offered to sample members located in the U.S. Any individuals who are known to work for the National Science Foundation will also be excluded from the incentive offer.

Additionally, NCSES will include an incentive experiment during the early-stage of data collection in 2023 to assess alternative methods of incentive delivery. One method uses a \$30 personalized check mailed to the sample member with the survey request, and the other uses a \$30 electronic VISA gift card delivered immediately upon survey completion. This is a change from the 2021 round, which primarily used prepaid debit cards. Offering the incentive prior to response with a personalized checks or offering contingent upon survey completion via an electronic gift card has the advantage over prepaid, non-personalized debit cards in that these incentive-offer mechanisms ensure incentives are only received by the targeted sample member.¹⁵ The use of electronic gift cards has a potential to shorten the lead time for incentive administration and, as a result, allows for real-time adaptive design incentive interventions that are not possible when using personalized check. The experiment will also evaluate the administrative cost, the time to complete, and number of contacting attempts required.

For details about the planned use of incentives in the 2023 SDR, refer to Supporting Statement Part B, Section 3.4.

¹⁴ <http://nap.napnationalacademies.org/26173>.

¹⁵ SDR is a mobile population and despite active attempts to locate current contact information, it is possible to have the wrong mailing address on file for individuals. Sending non-personalized incentives such as a non-personalized debit card or cash can result in the wrong individual receiving and using the incentive. Whereas a personalized check might be delivered to the wrong address, it is very unlikely to be cashed.

Incentive Costs

The proposed 2023 SDR incentive plan for the early and late phases to target an incentive offer to 25% of the total sample (n=31,500 cases) is summarized in the Table 1. According to this plan, a \$30 incentive will be offered to about 20,000 nonrespondents during the early stage in the 2023 SDR. In the late stage, a set of 11,500 nonresponding sample members who did not receive an early-stage incentive will be offered \$30.

In past SDR cycles where a check was mailed to sample members, a large number of the incentivized sample members did not cash the check despite responding to the survey. For example, in the 2017 SDR, 24,085 sample members were sent an incentive check. Of these individuals, only 10,078 cashed the incentive check (41.8%) and yet 15,114 of them completed the survey (62.8%). In 2017, very few individuals cashed their check and did not do the survey whereas a sizeable number did not cash their check yet did complete the survey. The same general pattern occurred in 2019 SDR with more of the incentivized sample members participating (56.2%) than the number who cashed their incentive (46.7%). However, in 2019, a higher proportion cashed the incentive check than in prior cycles and a smaller proportion of those incentivized responded to the survey. For 2023, we anticipate we will see similar patterns of survey cooperation with check cashing. In the group that receives the incentive only upon participation, we may see some participants deciding not to accept the virtual gift card at the conclusion of the survey, however, once a participant decides to accept the gift code, the cost of that gift card is fully incurred by the government whether the participant spends the gift card or not. In the 2023 SDR, the incentive plan (including the experiment) aims to offer a monetary incentive to 31,500 sample members with an estimated cost of \$476,850 with an estimated 15,895 individuals cashing the incentive offer and 17,640 individuals completing the survey.

Table 1: 2023 SDR Incentive Design by Data Collection Stage, Sample Type, and Incentive Type

Incentive Phase	Incentive Eligible Sample Type	Incentive Offer Field Period Week	Estimated Incentive Eligible Sample Size	Part of Experiment Yes/No	Incentive Groups		
					Check	Post-completion gift card	No Incentive
Early	Returning sample - only those who completed the 2019 and 2021 SDR after receiving and cashing the incentive	Week 3	4,500	No	4,500	NA	0
	New cohort - SED nonrespondents/soft refusals	Week 3	500	No	500	NA	0
	Returning sample- week 7 nonrespondents	Week 7	60,000	Yes	5,000	5,000	50,000
	New cohort - week 7 nonrespondents	Week 7	6,500	Yes	2,500	2,500	1,500
Late	Nonrespondents at the start of late-stage phase	Week 13	35,000	No	11,500	NA	23,500

10. ASSURANCE OF CONFIDENTIALITY

NCSES and its contractors are fully committed to protecting the confidentiality of all survey respondents. The SDR data will be collected under the authority of the National Science Foundation

Act of 1950, as amended, the America COMPETES Reauthorization Act of 2010, and the Confidential Information Protection and Statistical Efficiency Act (CIPSEA) of 2018. With the Federal Cybersecurity Enhancement Act of 2015, all respondent data are protected from cybersecurity risks through screening of the Federal systems that transmit the data. Cover letters (see Appendix E.1) and survey questionnaires (see Appendix D.2) to each selected sample member will advise them that the information they provide is confidential. The same notice of confidentiality will be used in the introduction to the CATI interview and will be displayed prior to the start of the survey in the online instrument. In addition, the CATI interviewers and online frequently asked questions (FAQs) will inform the respondents of the voluntary nature of their response and that the data will be used for statistical purposes only.

Standard data collection procedures incorporate numerous safeguards for protecting the data and must conform to a detailed security plan approved by NCSES. While collecting the SDR data, the information that could identify a sample member is separated from data about that person. Each sample member is assigned a unique identifier, and this identifier is used to store identifying information (such as name, address, etc.) in a separate, secure database apart from the survey response database. The SDR contractors and NCSES staff receive annual CIPSEA training to reinforce their legal obligations to protect the privacy and confidentiality of the SDR data; staff must sign data use agreements annually to acknowledge this legal obligation.

Completed SDR hard copy questionnaires and other contact materials will be housed in a secure storage room at the contractor's production facility. Only authorized staff – and only when necessary for data collection activities – will have access to hard copy materials from the SDR file room. Once information from the paper forms is captured and processing of the 2023 data is complete, the paper forms will be electronically imaged, and the paper forms will be securely destroyed. The contractor's electronic systems protect data via encryption at rest on its drive arrays, utilizing AES 256-bit encryption. All contractor systems for storage of electronic survey data will be protected by passwords available only to authorized study staff. For transmission of all documents and data, files are encrypted following the FIPS 140-2 compliance and transfer of all data utilizes the NIST 800-53 security framework.

The contractor will implement systems to make certain that data collected via the online questionnaire are secure. First, access to the online instrument will be allowed only with a valid Personal Identification Number (PIN) provided in the outreach material or verbally provided by a locator or telephone interviewer. Second, data transmission uses Secure Socket Layer (SSL) technology which establishes a secure and encrypted connection between the respondent's computer and the server on which the survey resides. If a respondent keeps an online survey open without any activity, the online server will close the connection after 7 minutes of inactivity, both preserving the data up to the break-off point and preventing unauthorized persons from completing the questionnaire. The front end of the web application will be displayed within a secured firewall on the demilitarized zone (DMZ) as an added layer of protection to the external environment. Response data, however, will reside inside the corporate domain past a second level firewall, which inspects all data transmissions entering or leaving the internal network by transferring responses from each web page as soon as respondents progress to the next web questionnaire page or screen. Moreover, the designated data storage location on the Local Area Network can only be accessed by the application on the DMZ through secure network connections, making it virtually impossible to connect with the internal network from the Internet. Both development and production servers will be backed up nightly as required by the contractor's disaster recovery plan.

NCSES and its contractors will analyze and make available SDR tabulations only in aggregate form and will take all measures necessary to assure that the identity of individuals or organizations will not be disclosed in either its statistical tabulations or in the SDR public-use micro-data files.

11. JUSTIFICATION FOR ASKING SENSITIVE QUESTIONS

No questions of a sensitive nature are asked in this data collection.

12. BURDEN ESTIMATE

For the 2023 SDR, a statistical sample of 125,246 individuals with U.S.-earned doctorates in science, engineering, or health will be contacted for participation in the survey. As with prior SDR data collection cycles, the sample consists of all eligible cases from the previous cycle (115,246), as well as a sample of 10,000 new PhD graduates. For 2023, the new cohort received their PhD between July 1, 2019, and June 30, 2021. Across the full production sample, approximately 111,904 individuals will reside in the U.S. and 13,342 will reside abroad. The amount of time to complete the questionnaire may vary depending on an individual's circumstances; however, on average, it will take approximately 22 minutes to complete the survey based on the response time from 2021 and changes to the survey instrument including the addition of the new retirement module. The 2015, 2017, 2019, and 2021 SDR data collections each resulted in an overall response rate of just under 70%. Assuming a 70% response rate (87,672 respondents) in 2023, the total survey burden for the 2023 SDR production sample is estimated to be 32,146 hours as shown in Table 2.

NCSES will initiate a program for between-cycle locating efforts to help reduce the effort and cost associated with locating sample members at the start of each cycle. The between-cycle period will occur several months before the start of pre-field locating. At this time, NCSES will attempt to contact approximately 30% of the production sample ($30\% \times 125,246 = 37,574$) to update or confirm their mail, email, and phone information. Of the approximately 30% contacted, we anticipate a 30% response rate based on a feasibility study conducted in late 2020 ($30\% \times 37,574 = 11,272$). In that study, it took on average 3 minutes to complete the request, for an estimated burden of 564 hours as shown in Table 2.

The overall burden hours for these 2 components of the SDR samples is 32,710 hours (32,146 for the survey plus 564 hours for between-cycle locating). Annualized over the requested clearance period of three years, the annual burden of the production survey is estimated to be 10,903 hours.

Table 2: 2023 SDR Estimated Burden Hours

SDR Collection	Sample Size	Respondents	Average Burden per Response	Overall Burden Hours
2023 SDR	125,246	87,672	22 minutes	32,146
Between Round Locating Update (between 2023 and 2025)	37,574	11,272	3 minutes	564
Total				32,710
Annualized Total over 3 years				10,903

The total cost to respondents for the 32,710 burden hours is estimated to be \$1,824,237. This is based on an estimated median annual salary of \$116,000 per full time employed SDR respondent from the 2021 data. Assuming a 40-hour workweek over 52-weeks of employment, this annual salary corresponds to an hourly rate of \$55.77. Over the three-year OMB clearance period, the average annual cost to the public for the 2023 SDR is estimated to be \$608,079.

13. COST TO RESPONDENTS

Not applicable. This survey will not require respondents to purchase equipment, software, or to contract out services.

14. COST TO THE FEDERAL GOVERNMENT

The total estimated cost to the Government for the 2023 SDR is approximately \$12.0 million, which includes survey cycle costs, and NCSES staff costs to provide oversight and coordination with the NSCG. The estimate for survey cycle costs is approximately \$11.5 million, which is based on sample size; length of questionnaire; CATI and online data collection technology; administrative overhead, design, printing, mail and telephone data collection costs; incentive payments; critical items data retrieval; data capture and editing; data quality control; imputation for missing item responses; weighting and estimating sampling error; file preparation and delivery of both cross-sectional and longitudinal files; preparation of documentation and final reports; analysis, and tabulations. The NCSES staff costs are estimated at \$562,500 (based on \$150,000 annual salary of 1.5 FTE for 2.5 years). Over the three-year OMB clearance period, the average annual cost to the Government is estimated to be \$4.0 million.

15. CHANGES FROM PRIOR CYCLE

The increase in burden of 63 hours from the 2021 SDR estimate of 32,647 hours to 32,710 hours expected in 2023 is a result of not including a bridge panel for 2023 which dropped the burden and offset the increased survey administration time for the new retirement module.

16. PLAN FOR TABULATION OR PUBLICATION

In general, the SDR data are presented as cross-tabulations of the data in reports and other data releases and continue to support time series indicators. The time schedule for the 2023 data collection and publication is currently estimated as follows:

Data Collection (Mail, CATI, online)	August 2023 – February 2024
Coding and Data Editing	September 2023 – March 2024
Final Edited/Weighted/Imputed Data File	June 2024
SDR InfoBrief	June 2024
SDR Detailed Statistical Tables	August 2024
SDR Public Use File	August 2024

17. DISPLAY OF OMB EXPIRATION DATE

The OMB Expiration Date will be displayed on the 2023 SDR questionnaire; in the online survey version, it will be included on the informed consent page of the online survey and available in the online FAQs; in the telephone interview, it will be read to sample members during the introductory informed consent. It

will also be included in the letterhead used to send sample members mailed correspondence and in emails to sample members.

18. EXCEPTION TO THE CERTIFICATION STATEMENT

Not Applicable. No exceptions to the certification statement are being sought.